

AI를 활용한 스마트 홈 서비스 연구

김지원¹, 박재영², 신민서³, 이진규⁴, 전지연⁵, 김현⁶

경상국립대학교 정보통계학과
고려대학교 건축사회환경공학부
울 로보틱스 기술연구소

zz0622@gnu.ac.kr, wodud2970@gmail.com, minseo3280@naver.com,
jink.lee0070@gmail.com, altmdbjfem32@gmail.com, kimhyun@khu.ac.kr

A Study on Smart home solution using AI

Ji-won Kim¹, Jae-young Park², Min-seo Shin³,
Jin-kyu Lee⁴, Ji-Yeon Jeon⁵, Hyun Kim⁶

¹Dept. of Statistics, Gyeong-sang National University
Environmental and Architectural Engineering, Korea University
Yul Robotics Technology Research Institute

요 약

본 연구는 'AI를 활용한 스마트 홈 서비스' 개발에 관한 것이다. 기존의 다양한 정보를 수집하고 제어하는 홈 IoT서비스에서 본 논문은 더 나아가, AI 기술을 바탕으로 사용자가 자신에게 맞는 형태로 스마트 디바이스들을 제어하여 홈 (Home)을 편리하게 제어할 수 있게 함과 동시에 사용자의 관여 없이도 AI를 활용해 자동으로 홈의 상황을 인지하고 동작할 수 있는 P2M + M2M 기술 기반 홈 IoT 서비스를 구현하고자 하였다. 특히 사용자의 동작을 인식해 IoT 기반 기기들을 통합적으로 제어할 수 있도록 모션인식, 영상인식 기술 등 사용자를 인식하고 주변 환경 상태를 실시간으로 측정해 최적의 제어 서비스를 제공하는 것을 목적으로 하였다.

1. 서 론

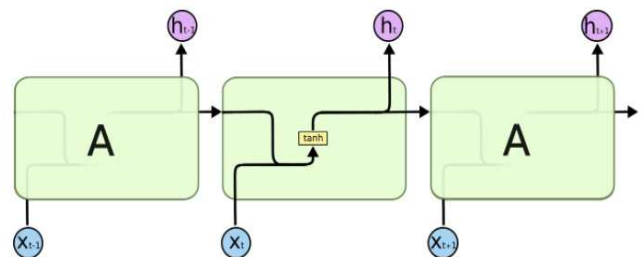
최근 1인 가구의 지속적인 증가와 코로나 여파 등으로 인해 사람들이 집에 머무는 시간이 증가하고 있다. 이에 따라 효율적으로 집을 관리할 수 있는 IoT 기술이 탑재된 가전제품의 판매량 역시 증가하고 있는 추세이다. 우리 연구팀은 높아지고 있는 IoT 스마트 홈 디바이스 관심도에 기인해 IoT 기술을 통하여 집을 효율적으로 관리하는 것에 주목하였다. 그리고 AI를 활용한 사용자의 행동 인식을 바탕으로 기기를 작동하게 만들어 사용자 편의성을 크게 증가시킬 수 있는 다양한 서비스들을 개발하고자 하였다. AI를 바탕으로 IoT 기기들을 최적 상태로 관제할 수 있도록 하고자 하였다.

2. 관련연구

2.1 LSTM

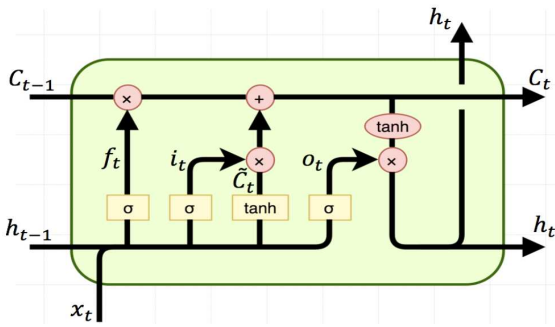
순환신경망은 기존의 일반 신경망 모델에서 시계열 개념이 추가된 딥러닝 모델로 일반신경망 모델과 달리 지난 시점의 정보를 기억해 다음 시점의 데이터에 어떤 영향을 주는지 학습할 수 있고, 이러한 장점으로 인해 순환신경망은 입력이 순차적으로 이뤄지는 시계열 데이터 예측에 큰 효율을 발휘한다.

순환신경망의 구조는 그림 1과 같다.



<그림1. RNN구조>

본 연구에서 사용된 LSTM은 기울기가 사라지는 문제를 극복하기 위해서 제안된 모델로, LSTM은 그림 2와 같이 전체 체인을 관통하는 셀 상태를 통해 과거 학습결과를 큰 변화 없이 전달하는 구조로 장기 의존성 문제를 해결한다. LSTM에서 입력 게이트는 입력 값을 얼마나 받아들일지, 망각 게이트는 어느 정도 정보를 잃어버릴지, 출력게이트는 얼마나 출력할지 파라미터를 찾아나가며 가장 예측을 잘할 수 있는 방향으로 적절하게 이용된다.



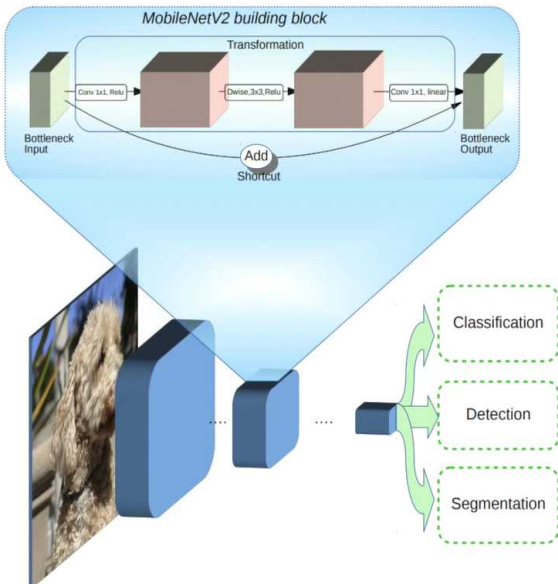
<그림2. LSTM구조>

$$\begin{aligned}
 f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \\
 i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \\
 o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \\
 c_t &= f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \\
 h_t &= o_t \circ \sigma_h(c_t)
 \end{aligned}$$

<그림3. LSTM수식>

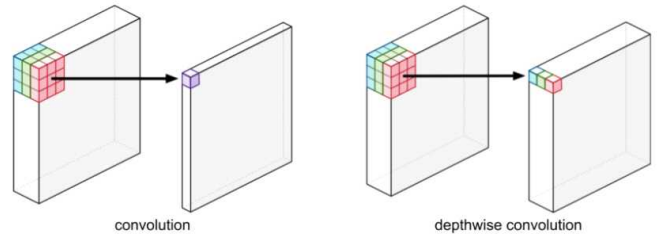
2.2 MobileNetV2

MobileNetV2은 모바일이나 임베디드에서도 real-time으로 작동할 수 있게 모델을 경량화하고, 성능 또한 유지한 네트워크이다.



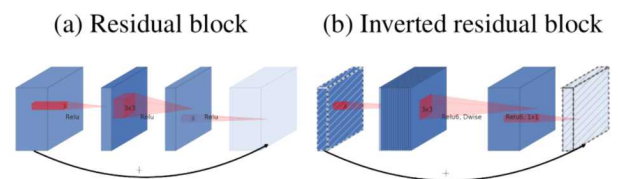
<그림4. MobileNetV2 구조>

MobileNetV에서는 그림5와 같이 depthwise separable convolution을 사용해 크기를 줄이면서 표현력을 강화시켜 연산량을 획기적으로 감소시킨다.



<그림5. Depthwise convolution>

또한, Inverted residual block을 사용하여 기존의 residual block과는 반대로 채널을 축소시키고 Depthwise convolution을 하는 것이 아니라 채널을 확장하고 Depthwise convolution을 하여 정보의 손실을 최소화 하였습니다.

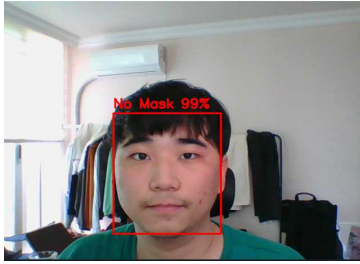


<그림 6. Residual block , Inverted residual block>

본 연구 내 마스크 착용 인식에 MobileNetV2를 사용하였다. MobileNetV2는 서버에서 밝혔듯이 연산량이 많지않아 임베디드에서도 real-time으로 작동이 가능하고 표현력 또한 다른 경량화 모델에 비해 성능이 높기 때문에 예측 모델로 선정하게 되었다.

2.3 영상인식을 통한 사용자 분석

코로나로 인해 외출 시, 무조건 마스크를 착용해야 하기 때문에 영상인식 기술로 마스크 착용여부를 빠르게 판단하여 사용자에게 알릴 수 있는 서비스를 구현하였다. 마스크 인식을 위해 사람의 얼굴을 예측 한 후, 마스크 사용유무 예측을 진행해야 하므로 딥러닝 기반으로 제작된 Dlib의 Face_recognition 라이브러리를 사용하여 얼굴인식을 하였고, 계산복잡도와 사이즈를 줄이면서 Conv층의 표현력을 유지하기 위해 Depthwise separable convolution과 Linear bottlenecks 현상을 방지 시켜주는 main fold방식을 사용하였다. Mobilenetv2와 Dense층을 결합해 마스크 인식 모델을 690개의 마스크를 착용한 사용자 데이터와 686개의 마스크를 착용하지 않은 사람의 데이터를 통해 모델을 학습시켜 예측하여 그림 7의 결과 값을 획득할 수 있었다.



<그림 7. 마스크 착용 여부 판단 예시>

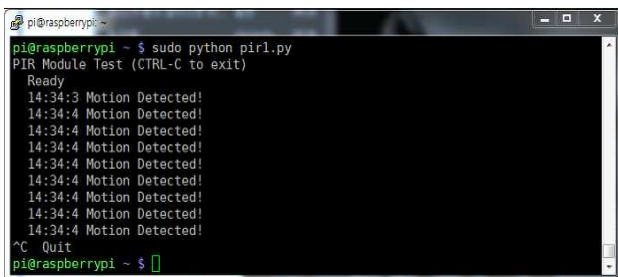
2.4 모션 제어를 통한 스마트 디바이스 제어

음성인식을 기반으로 하는 제어 서비스는 음성을 통한 '호출 -> 음성 전달-> 서버 -> 결과 회신'로 처리하는 과정에서 짧게는 5초에서 길게는 10초 이상 시간이 소요된다. 우리는 이러한 음성 제어를 대신해 모션으로 빠른 제어가 가능함을 확인할 수 있었고, 이를 기반으로 하는 서비스를 구현하고자 하였다. 모션인식을 통한 제어는 약 -30frame의 시점까지 사람의 동작을 예측해 제어를 하기 때문에 음성인식 서비스 보다 제어시간을 단축시킬 수 있는 강점을 지니고 있다. 우리는 시계열 데이터인 행동에 대한 패턴을 학습하고 행동 예측에 적합한 모델을 구현하였다.



<그림 8. 모션 인식 제어 예시>

특정 모션에 따라 제어를 할 수 있도록 하고자 하였는데 우선, 예측을 위해 박수를 치거나, 손을 드는 등 특정한 행동에 대한 학습을 위해 python package중 mediapipe의 landmark detection Algorithm을 통해 33개의 몸체 landmark 와 40개의 손 landmark를 통해 각 landmark 당 x, y, z, v(visualiable) 4가지 정보에 대해 1 frame 당 총 292개의 정보를 기반으로 구현하였다.



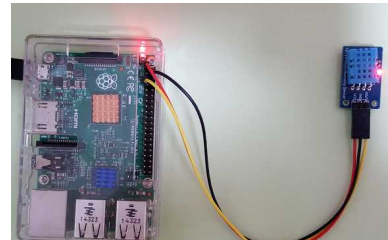
<그림 9. 모션 인식 값 출력>

시작 프레임부터 지난 30frame까지의 landmark 데이터를 하나의 sequence로 정의해 하나의 동작으로 묶이게 설계하였다. 그림3은 몸과 손의 landmark를 시각화 한 것이고 그림 4는 모션을 인식 여부를 보여주고 있다.

서두에 언급했듯이 30frame 단위로 시계열 예측을 해야 했기에 딥러닝 모델 중 순환신경망인 LSTM과 FC layer를 결합한 모델을 개발했다.

2.5 IoT센서를 통한 자동 제어

우리는 조도 센서와 미세먼지 센서, 빗물 센서를 통해 외부 대기 상황을 실시간으로 측정하여 미세먼지 높거나 비가 온다고 판단될 경우, 자동적으로 창문을 닫아주는 서비스를 구상하였다. 사용자의 직접적인 조작 없이 제어가 가능해 사용자 편의성을 높일 수 있다는 특징을 갖고 있다. 또한, 압력 센서를 바탕으로 사용자의 취침과 기상을 인식해 커피포트 발열, 커튼 개폐와 같은 사용자에게 맞는 서비스를 제공한다.



<그림 10. IoT센서>

3. 결론

본 연구에서는 기존의 IoT기기를 제어하는 방식인 음성 인식과 수동 인식에서 벗어나 다양한 방법으로 기기를 제어 할 수 있는 아이디어를 제안하였다. 제안한 방법으로는 모션인식과, 마스크 착용인식, IoT센서를 통한 자동인식 등이 있다. 모션인식은 LSTM layer와 Dense layer를 통해 feature들의 특징을 뽑고 예측하는 과정을 통해 사용자의 모션을 인식하는 방법을 제안했고, 마스크 착용 인식은 Face Recognition 라이브러리를 통해 마스크 착용 분별 부분의 영역을 예측하고 MobileNetV2를 통해 최종적으로 마스크 착용 여부를 판단하는 알고리즘을 구현했고, IoT센서를 활용한 자동 제어는 압력, 조도, 미세먼지, 빗물 센서 등 다양한 센서들의 값을 받아 자동적으로 제어가 가능한 프로그램 설계를 하였다. 이러한 과정을 통해 우리는 기존보다 더 효율적이게 디바이스를 제어함으로써 홈 IoT기반 기기들을 사용함에 있어서 사용자에게 더 나은 UX를 전달하였다. 추 후 과제로는 본 연구에서 사용되었던 동작예측 모델이나, mobileNetV2를 사용하기보다 카메라에서 이미지를 받아 서버로 전달하는 과정을 통해 많은 연산량을 감당할 수 있는 GPU를 사용해 기존보다 정확도를 높이는 방법을 개발할 필요가 있다.

3.1 기대효과

[1] 다양한 정보를 수집하고 제어하는 기존에 서비스에서 더 나아가, 사물이 스스로 학습하여 자동으로 인간에게 최적화 된 환경을 제공한다면 이는 인간에게 지금까지와는 차원이 다른 라이프 스타일을 기대할 수 있다.

[2] 임베디드 + 모바일 + 클라우드에 개인 맞춤형 서비스 까지 추가된 새로운 형태의 IoT기반 제품을 개발하여 사용자에게 매력적이고 차별성을 갖춘 획기적인 서비스를 제공해주어 편의성을 제공해 줄 수 있다.

3.2 활용

[1] 사용자의 개입을 최소화 하면서도 최적의 환경을 제공해준다. 또한, 개인에 맞추어진 개인 맞춤형 서비스로 최적의 환경을 제공해주어 홈 서비스를 이용하는 사용자들에게 차별화되어있어 매력적인 IoT 스마트 홈 서비스를 사용하게 될 것이다.

[2] AI 기술을 활용한 동작 인식은 제어기기를 소지하고 있지 않더라도 집안 내 어디서든 제어가 편리한 음성인식의 장점과 음성인식보다 빠르게 제어할 수 있는 앱을 통한 제어의 장점을 결합해 IoT 기기를 가장 쉽게 제어할 수 있는 방법이 될 것이다.

[3] 서로 다른 기종들끼리도 손쉽게 자동으로 제어가 가능한 새로운 기술로 활용가능하다.

Conference on Computer Vision and Pattern Recognition

[6]HAKIM, Noorkholis Luthfil, et al. Dynamic hand gesture recognition using 3DCNN and LSTM with FSM context-aware model. Sensors, 2019

※ 본 프로젝트는 과학기술정보통신부 정보통신창의인재양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.

참고문헌

[1] Roger S. Pressman "Software Engineering A Practitiners' Approach" 3rd Ed. McGraw Hill

[2] kairess [2020] [COVID-19 Face Mask Detection]
<https://github.com/kairess/mask-detection>

[3] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition

[4] 강승우, "라즈베리 파이에서 MQTT 프로토콜 이용 시 전력 소모 분석" Journal of the Korea Institute of Information and Communication Engineering 21(12), 2017.12,

[5] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov and Liang-Chieh Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE